

# Twenty years of RNA crystallography

ERIC WESTHOF

Architecture et Réactivité de l'ARN, Institut de Biologie Moléculaire et Cellulaire du CNRS, Université de Strasbourg, 67084 Strasbourg Cedex, France

I still remember vividly a phone call in the spring of 1995. I was then modeling an RNA structure in the darkened room where the Evans & Sutherland PS300 was buzzing loudly. Tom Cech was on the line and, after the usual greetings, he asked with his usual jovial and direct way, "Eric, we want to start a new journal dedicated to RNA; are you willing to become a journal editor?" I was stunned and speechless. My knowledge of RNA biology was scanty. I knew quite well all the twenty or so RNA structures that existed back then, as I had used their structural information to build RNA models in three dimensions, such as the one I was building when Tom called, or that of the core of group I introns I had built with François Michel a few years earlier. I remember that particular model to be quite appreciated by the community for rationalizing experimental data and helping to devise new experiments. But I did not think I qualified to become a journal editor! Tom insisted, and answered my novice questions about the job of an editor. Excited but still not fully realizing what was happening, I accepted and never regretted it.

Tom's insightful initiative followed a logical move, since several RNA leaders had established the RNA Society in 1993 and organized the first RNA meeting, as we now know them, in Madison, Wisconsin, at the end of May 1996. The *RNA* journal, hence also supported by the RNA Society, became a critical avenue for publishing the exciting research that had been shared first as posters or talks at the popular RNA meetings. Over the years, the RNA meetings, always extremely well attended, took place in various places around the world with the lively poster sessions being always a fantastic playground for meeting authors and enriching one's list of potential referees. The poster rooms were rustling about RNAs at the origin of life, sequence alignments, dynamics of nucleic acids, magnesium ions, RNA evolution and folding, RNA purification and stability, novel RNAs being discovered in cells. Many of those posters appeared later as articles in the *RNA* journal. During the oral sessions, one could learn anything that was new in RNA biology in general, splicing, nonsense mediated decay, RNA editing or maturation. In this short note, realizing the amazing expansion of

RNA biology, I would like to focus on the developments of RNA structural biology.

In 1995, the number of X-ray RNA structures (alone or in complex with protein) in the Protein Data Bank amounted to about 1% of the present content of total RNA and RNP structures. One year later, a breakthrough in RNA crystallography happened with the crystal structure of the P4-P6 domain of the *Tetrahymena* group I intron, a structure twice the size of a tRNA with novel types of RNA contacts, from the labs of Tom Cech and Jennifer Doudna. Thus, it was a critical moment in the overall perception of the central role of RNA in molecular biology; and Tom was a very active and central figure in the push for promoting research in RNA chemistry, biochemistry, structure, and biological functions. With the beautiful structure of P4-P6, many interested scientists knew that large RNAs could also be crystallized starting from *in vitro* synthesis and production. And the repertoire of surprising and beautiful RNA architectures expanded at a very quick pace. Over the years, almost every single cover of the RNA Meeting abstract book and of the *RNA* journal displayed a novel crystal structure of an RNA or RNA/protein complex. The number of RNA and RNP structures in the PDB grew from the 31 available in 1995 to about five times more in 2000 (155 structures), to 15 times in 2005 (462 structures) and, at the beginning of 2015, to close to 60 times the number of RNA and RNP structures (2495) 20 years earlier. Certainly, not all those structures are independent structures; several are variants in sequence or ligand binding. However, about half of the total number of structures is at a resolution worse than 3.3 Å.

The late 1990s and early 2000s were the big days of ribozymes; and group I introns came to epitomize catalytic RNAs. Surprisingly, however, after the P4-P6 structure, structures of group I introns from various species appeared at a slow pace, starting with *Tetrahymena* in 1998 and 2003, *Azoarcus* in 2004, and the phage *Twort* in 2005. More recently (2014), the structure of the lariat capping ribozyme was also solved. Despite several structures of nucleolytic ribozymes (hammerhead [1994], hepatitis delta virus [1998], hairpin [2001], and recently the twister [2014]), the understanding

Corresponding author: [e.westhof@ibmc-cnrs.unistra.fr](mailto:e.westhof@ibmc-cnrs.unistra.fr)

Article and publication date are at <http://www.rnajournal.org/cgi/doi/10.1261/rna.049726.115>. Freely available online through the *RNA* Open Access option.

© 2015 Westhof This article, published in *RNA*, is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

pathway was tortuous and it took 12 years (1994–2006) to obtain a fully biologically relevant structure of the hammerhead ribozyme. Several factors had to be deconvoluted for reaching some consensus in our chemical understanding; among others, the interplay between folding and catalysis, the distinction between structural and catalytic divalent ions, and the roles of bases through protonation and deprotonation in the chemical steps. Although it took only 10 years between the first structures of the bacterial 30S and 50S ribosomal particles and the structure of the eukaryotic ribosome, it took seven years between the first crystal structure of an RNaseP and of its holoenzyme. For the family of group II introns, since the first crystal structure in 2008, not much fundamentally novel has appeared at the structural level.

A field that yielded an amazingly rich chemically and architecturally diverse harvest is that of the riboswitches discovered in 2002. The PDB contains more than 100 structures of the aptamer domain of riboswitches, with and without ligands and under various conditions, with about 20 of them solved at a resolution better than 2.0 Å. Those riboswitch aptamer domains recognize a huge variety of ligands, from the tiny fluoride ion to the cyclic second messengers and the large adenosylcobalamin of vitamin B12, and offer striking examples of the potential of folded RNA molecules for molecular recognition. Thus, with only the four main bases and without any chemical modification, one could recognize almost any atomic or molecular entity. All those structures demonstrated that RNA could reach its full molecular recognition potential through convoluted and complex three-dimensional architectural folds exploiting all the stereochemical capabilities of the RNA backbone, recurrent RNA modules, and especially non-Watson–Crick base pairs. Clearly, everything that chemistry allows is at some stage exploited in biology. The recent demonstration that tautomerism of G and U bases contributes to the intrinsic lower limit of fidelity errors in translation upholds that view. The roles of the modified tRNA bases in the anticodon loop constitute now a major stream of research in cellular stress.

Thus, those first 20 years of RNA structural research taught us, among other things: (i) that all hydrogen bonding capabilities of the bases are exploited for forming the base pairs necessary for the beautifully intricate architectural folds that RNA can adopt to bind other RNAs, proteins, or encapsulate ligands, (ii) that protonation and deprotonation of

bases contribute to the catalysis of nucleolytic ribozymes, (iii) that tautomeric forms of bases lead to Watson–Crick-like base pairs that fool the translational fidelity mechanisms of ribosomes and polymerases.

Structural biology of RNA, especially with X-ray crystallography, made possible these advances in our understanding. A huge amount of effort put forth by several researchers went into the experimental and theoretical development of tools for deriving accurate and cogent crystallographic structures with biological relevance. The success of crystallography is undoubtedly due to the fact that it is firmly and solidly grounded in physics and chemistry. However, despite the numerous mathematical or chemical safeguards and warnings available, errors in stereochemistry, in RNA folds or in solvent or ligand identifications do occur and are, unfortunately, rarely corrected afterwards in the main depository databases that are constantly used by scientists all over the world for their research work.

RNA structural biology is nowadays an extremely complex field, arduous experimentally and theoretically, with previous structures being used in the integration of more complex assemblies in cryo-electron microscopy and in X-ray crystallography. In this respect, journal editors and referees have a major role to play in order to prevent incorrect structures to reach the literature and, almost worse, the depository databases.

In order to continue to maintain the highest level of usefulness and accuracy in the structural databases, I would like to urge referees (for the *RNA* journal and for any journal) to request complete and detailed X-ray statistics tables, validation reports and quality indicators, coordinates, and electron-density maps for proper evaluation and, as usual, to help the authors. Despite fierce competition for publication, I am sure authors will comply with such requests. The serious and fair peer-review process offers enough rail guards to prevent any mishandling of the X-ray data obtained after so much dedication and efforts by the authors. I feel it apt to conclude with words written by Alex Wlodawer and coworkers in 2013 (*FEBS J*), “As a community, we have to make sure that the high level of advancement is preserved and expanded, and also that the future generations of structural biologists are prepared to gain deeper insight from the massive amounts of data, and to take crystallography to areas that today may not even be foreseen.”



# RNA

A PUBLICATION OF THE RNA SOCIETY

## Twenty years of RNA crystallography

Eric Westhof

RNA 2015 21: 486-487

---

**Open Access**

Freely available online through the *RNA* Open Access option.

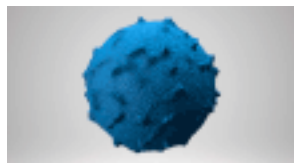
**Creative Commons License**

This article, published in *RNA*, is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

**Email Alerting Service**

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

---



Work with our RNA experts to  
find biomarkers in exosomes.

**EXIQON**

---

To subscribe to *RNA* go to:

<http://rnajournal.cshlp.org/subscriptions>

---